Introdução à Otimização

Workshop PADSW05

Magno Severino
Programa Avançado em Data Science e Decisão

3 de dezembro de 2022

Objetivos de aprendizagem

Ao final deste workshop você será capaz de

- Definir um problema de otimização;
- Modelar um problema de otimização;
- Desenvolver uma solução para um problema de otimização linear;
- Relacionar a otimização com os conceitos de aprendizagem estatística.

Agenda

1. O que é otimização matemática.
2. Características dos problemas.
3. Como expressá-los matematicamente.
4. Solução através do método Simplex.
5. Relação com aprendizagem estatistica.

Modelos matemáticos

$$\text{Lucro} = \text{Receita} - \text{Despesas}$$

$$\text{Lucro} = f(\text{Receita}, \text{Despesas})$$

$$Y = f(X_1, X_2)$$

Em particular,

$$f(X_1, X_2) = X_1 - X_2$$

Generalizando,

$$Y = f(X_1, X_2, \ldots, X_p)$$

Modelos Prescritivos *versus* Modelos Preditivos

Categoria	**Forma de** $f(\cdot)$
Modelos Preditivos	desconhecido, mal definido
Modelos Prescritivos	conhecida, bem definida

Introdução

- Nosso mundo possui recursos limitados.
    - A quantidade de petróleo que podemos retirar da terra é restrita.
    - As empresas tem um número limitado de trabalhadores.
    - Cada um de nós tem uma quantidade de tempo limitada para realizar ou desfrutar as atividades que programamos para todos os dias.
    - A maioria de nós tem uma quantidade limitada de dinheiro para gastar na realização dessas atividades.
- Objetivo: decidir como melhor usar os recursos limitados disponíveis a um indivíduo ou uma empresa.
- Problema: como alocar os recursos de maneira a maximizar os lucros ou minimizar os custos?
- A Programação Matemática é uma área da *business analytics* que encontra a maneira mais eficiente de usar recursos limitados para atingir os objetivos definidos.
- Por esse motivo, a programação matemática é geralmente chamada **otimização**.

Aplicações da otimização matemática

- Exemplo 1: **Determinação do Mix de Produtos**
  - Uma indústria produz diferentes produtos a base de leite.
  - Cada produto requer diferentes quantidades de matérias-primas e mão de obra.
  - De maneira semelhante, a quantidade de lucro gerado pelos produtos varia.
  - A gerência da empresa deve decidir *quanto* de cada produto produzir para *maximizar* os lucros ou *atender à demanda* com custo mínimo.

Aplicações da otimização matemática

- Exemplo 2: **Roteamento e Logística**
- Uma rede de lojas de eletrodomésticos têm armazéns em todo o país, os quais são responsáveis por manter as lojas abastecidas com mercadorias.
- A quantidade de mercadorias disponíveis nos armazéns e a quantidade necessária em cada loja tendem a flutuar.
- O custo da remessa e da entrega de mercadorias dos armazéns para os locais de varejo também varia.
- Grandes somas de dinheiro podem ser economizadas por meio da determinação do método mais barato de transferência de mercadorias dos armazéns para as lojas.

Características dos problemas de otimização

Os problemas de otimização envolvem três elementos:

- Tomar uma ou mais **decisões**:
  - Ex. 1: quanto de cada produto deverá ser produzido?
  - Ex. 2: quanto de cada produto deverá ser enviado de cada armazém para as diversas lojas?
- Respeitar **restrições** com relação às alternativas disponíveis.
  - Ex. 1: ao determinar o número de produtos a ser fabricado, um gerente de produção provavelmente enfrentará problemas com uma quantidade limitada de matérias-primas e de mão de obra.
  - Ex. 2: há uma limitação física com relação à quantidade de mercadoria que um caminhão pode carregar de um armazém para as lojas em sua rota.
  - Não é incomum que problemas de otimização do mundo real tenham centenas ou milhares de restrições.
- Existência de uma meta ou **objetivo**.
  - Ex. 1: escolher o mix de produtos que *maximizará* os lucros.
  - Ex. 2: identificar a rota que *minimizará* o custo total com o transporte.

Como expressar matematicamente

- Três elementos: decisões, restrições e um objetivo.
- Precisamos definir símbolos matemáticos para representar cada um desses três elementos.

Decisões

- Representadas por $X_1, X_2, \ldots, X_p$, chamadas de **variáveis de decisão**.
- Elas representam as quantidades de diferentes produtos que o gerente pode escolher produzir, por exemplo.

Como expressar matematicamente

- Três elementos: decisões, restrições e um objetivo.
- Precisamos definir símbolos matemáticos para representar cada um desses três elementos.

Restrições

- Função das variáveis de decisão que deve ser menor ou igual a, maior ou igual a, ou igual a um valor específico, por exemplo:

$$
\begin{aligned}
f_1(X_1, X_2, \cdots, X_p) &\leq b_1\,, \\
f_2(X_1, X_2, \cdots, X_p) &\geq b_2\,, \\
f_3(X_1, X_2, \cdots, X_p) &= b_3\,.
\end{aligned}
$$

Como expressar matematicamente

- Três elementos: decisões, restrições e um objetivo.
- Precisamos definir símbolos matemáticos para representar cada um desses três elementos.

Objetivo

- Identifica alguma função das variáveis de decisão que o tomador de decisão deseja **maximizar** ou **minimizar**.

Como expressar matematicamente

Formato geral de representação de um problema de otimização:

$$
\begin{aligned}
\text{MAX (ou MIN):} \quad & f_0(X_1, X_2, \ldots, X_p) \\
\text{Sujeito a:} \quad & f_1(X_1, X_2, \ldots, X_p) \leq b_1 \\
& \vdots \\
& f_k(X_1, X_2, \ldots, X_p) \geq b_k \\
& \vdots \\
& f_m(X_1, X_2, \ldots, X_p) = b_m
\end{aligned}
$$

Um exemplo de problema de Programação Linear

- Uma empresa fabrica e vende dois modelos de banheiras: a Aqua-Spa e a Hydro-Lux.
- O gerente precisa decidir quanto de cada tipo de banheira produzir em seu próximo ciclo de produção.
- Ele compra cubas de fibra de vidro pré-fabricadas de um fornecedor local e adiciona a elas a bomba e a tubulação para produzir as banheiras. (Esse fornecedor pode abastecê-lo com quantas cubas forem necessárias.)
- O mesmo tipo de bomba é instalado em ambos os modelos de banheira.
- Ele terá apenas 200 bombas disponíveis durante o próximo ciclo de produção.
- Do ponto de vista da fabricação, a principal diferença entre os dois modelos de banheira é a quantidade de tubulação e de trabalho necessários.
- Cada Aqua-Spa requer 9h de trabalho e 12m de tubulação. Cada Hydro-Lux requer 6h de trabalho e 16m de tubulação.
- Haverá 1566h de trabalho de produção e 2880m de tubulação disponíveis durante o próximo ciclo de produção.
- Ele tem lucro de $\$350$ em cada Aqua-Spa vendida e, em cada Hydro-Lux que comercializa, $\$300$.
- Ele está confiante de que poderá vender todas as banheiras que produzir.
- A pergunta é: quantas Aqua-Spas e Hydro-Luxes deve-se produzir com o objetivo de maximizar os lucros durante o próximo ciclo de produção?

Etapas na formulação de um modelo de PL

1. Entenda o problema.
2. Identifique as variáveis de decisão.
   - Responda: quais são as decisões fundamentais que devem ser tomadas para resolver o problema?
3. Coloque a função objetivo como uma combinação linear das variáveis de decisão.
4. Coloque as restrições como combinações lineares das variáveis de decisão.
5. Identifique quaisquer limites nas variáveis de decisão.
   - Normalmente, limites superiores ou inferiores simples são aplicados às variáveis de decisão.

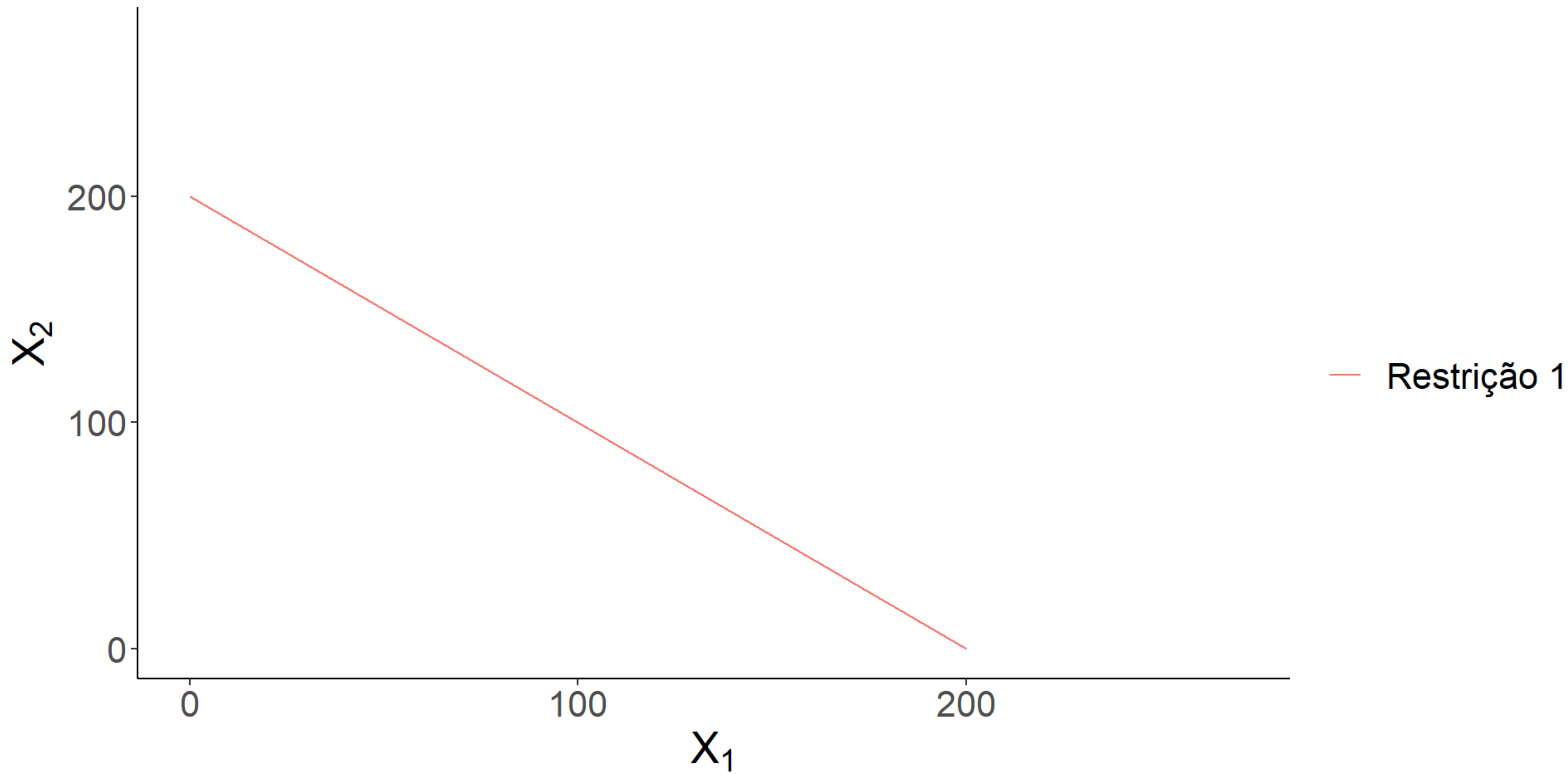
Solução

Formulação matemática

$$
\begin{aligned}
\text{MAX:} \quad & 350X_1 + 300X_2 \\
\text{Sujeito a:} \quad & 1X_1 + 1X_2 \leq 200 \\
& 9X_1 + 6X_2 \leq 1566 \\
& 12X_1 + 16X_2 \leq 2880 \\
& X_1 \geq 0 \\
& X_2 \geq 0
\end{aligned}
$$

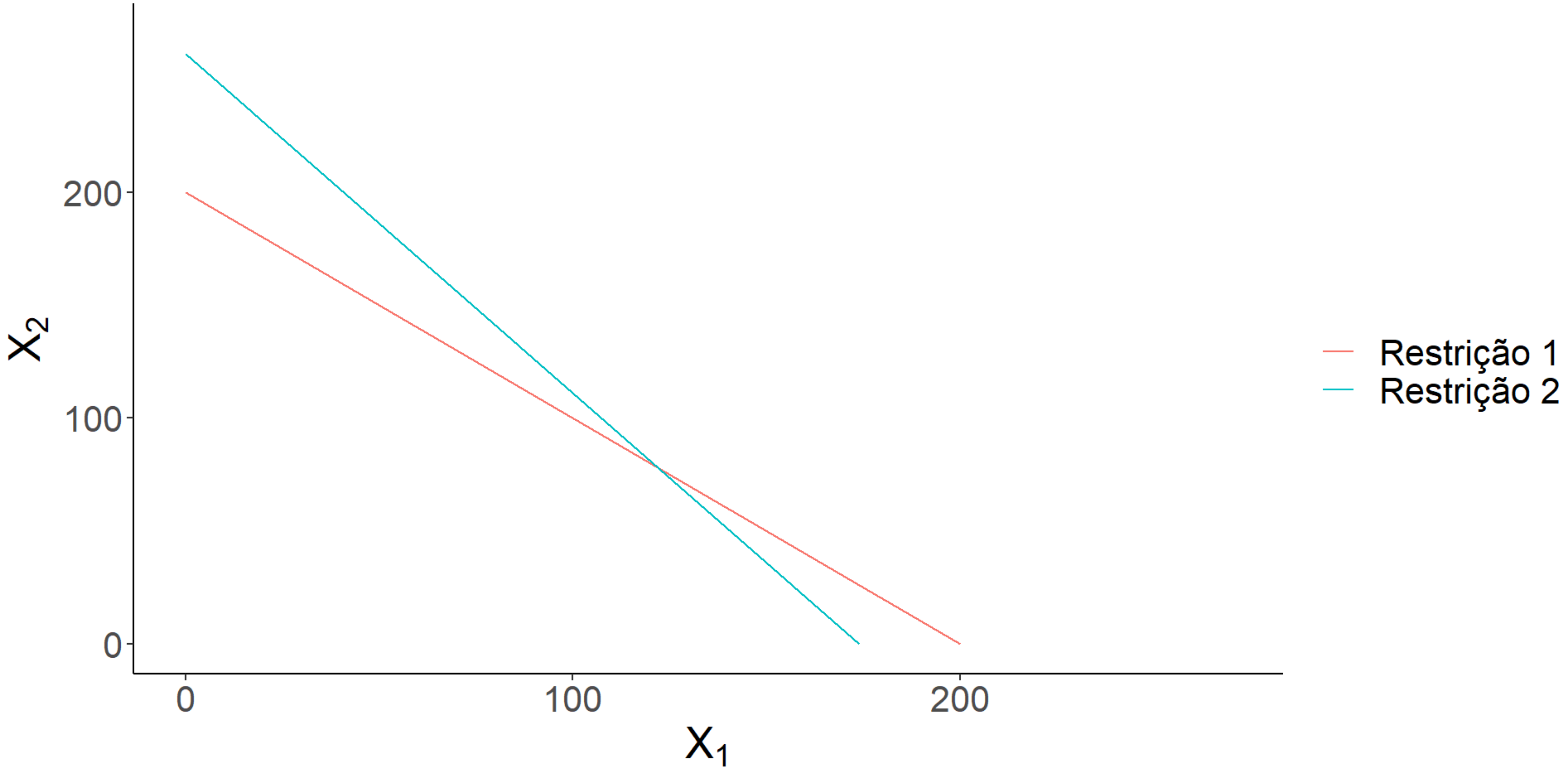
Encontrando a solução do problema

Restrição 1: $X_1 + X_2 = 200$



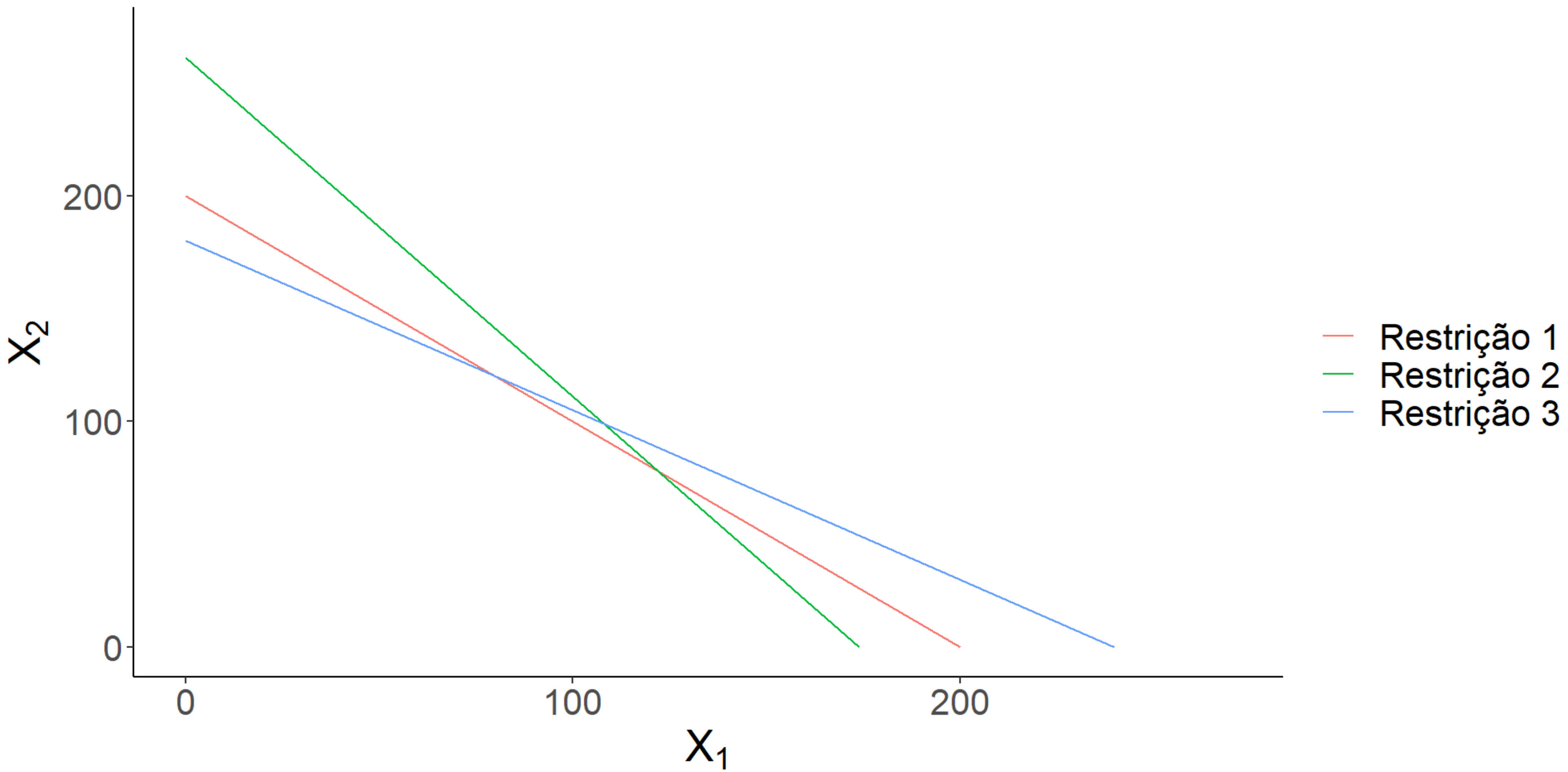
Encontrando a solução do problema

Restrição 2: $9X_1 + 6X_2 = 1566$



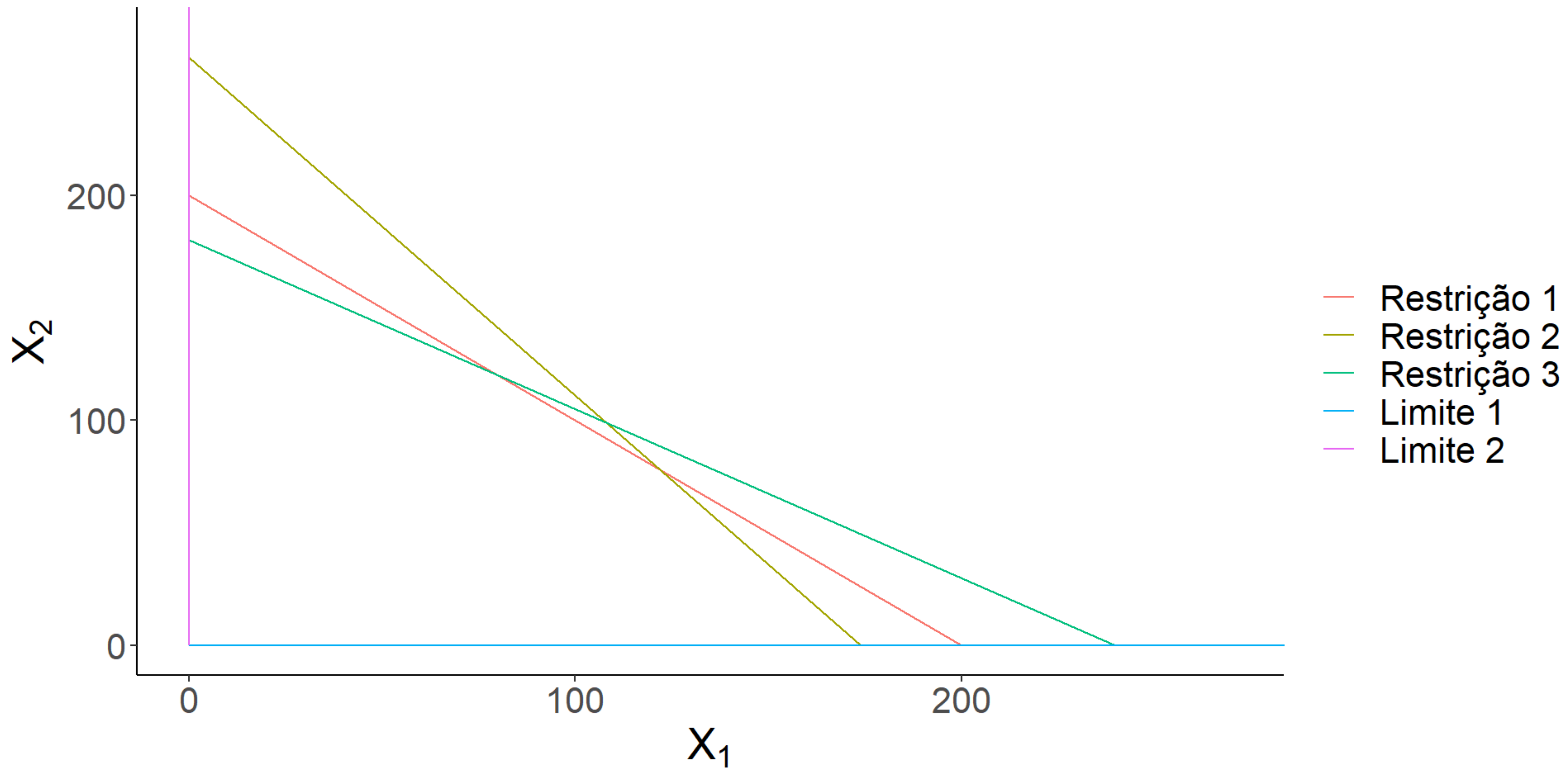
Encontrando a solução do problema

Restrição 3: $12X_1 + 16X_2 = 2880$



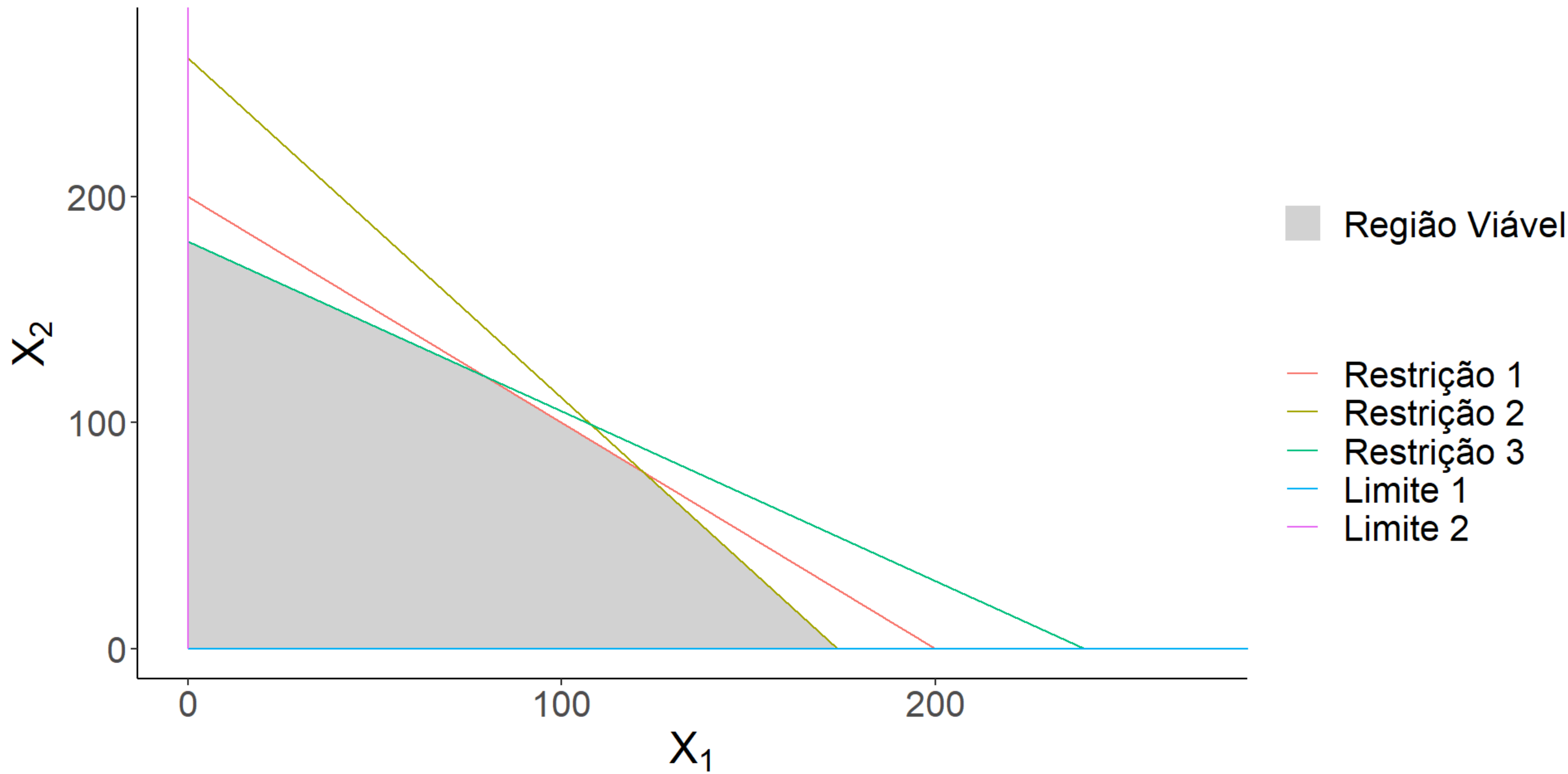
Encontrando a solução do problema

Limites: $X_1 \geq 0, X_2 \geq 0$



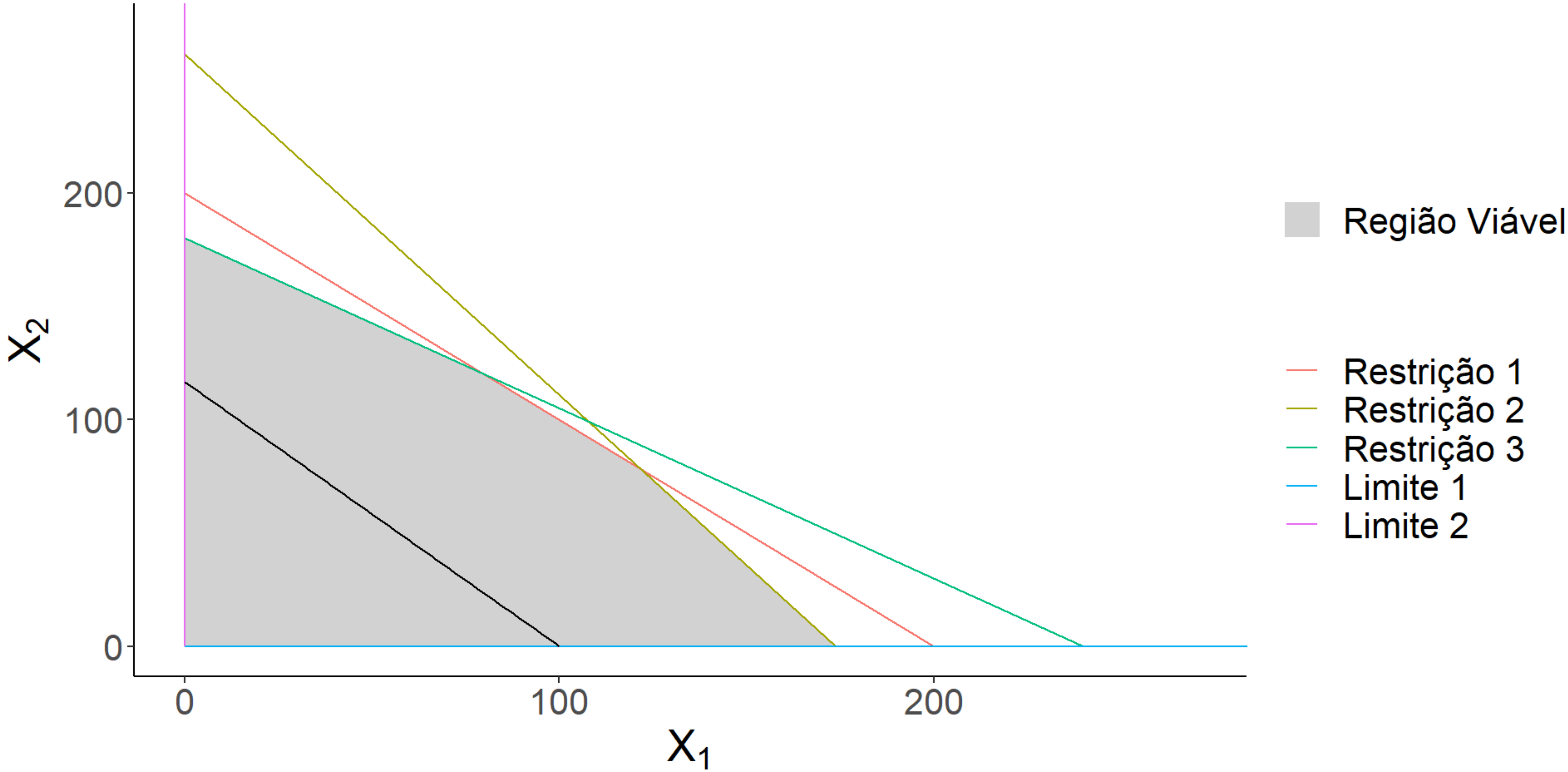
Encontrando a solução do problema

Região de viabilidade



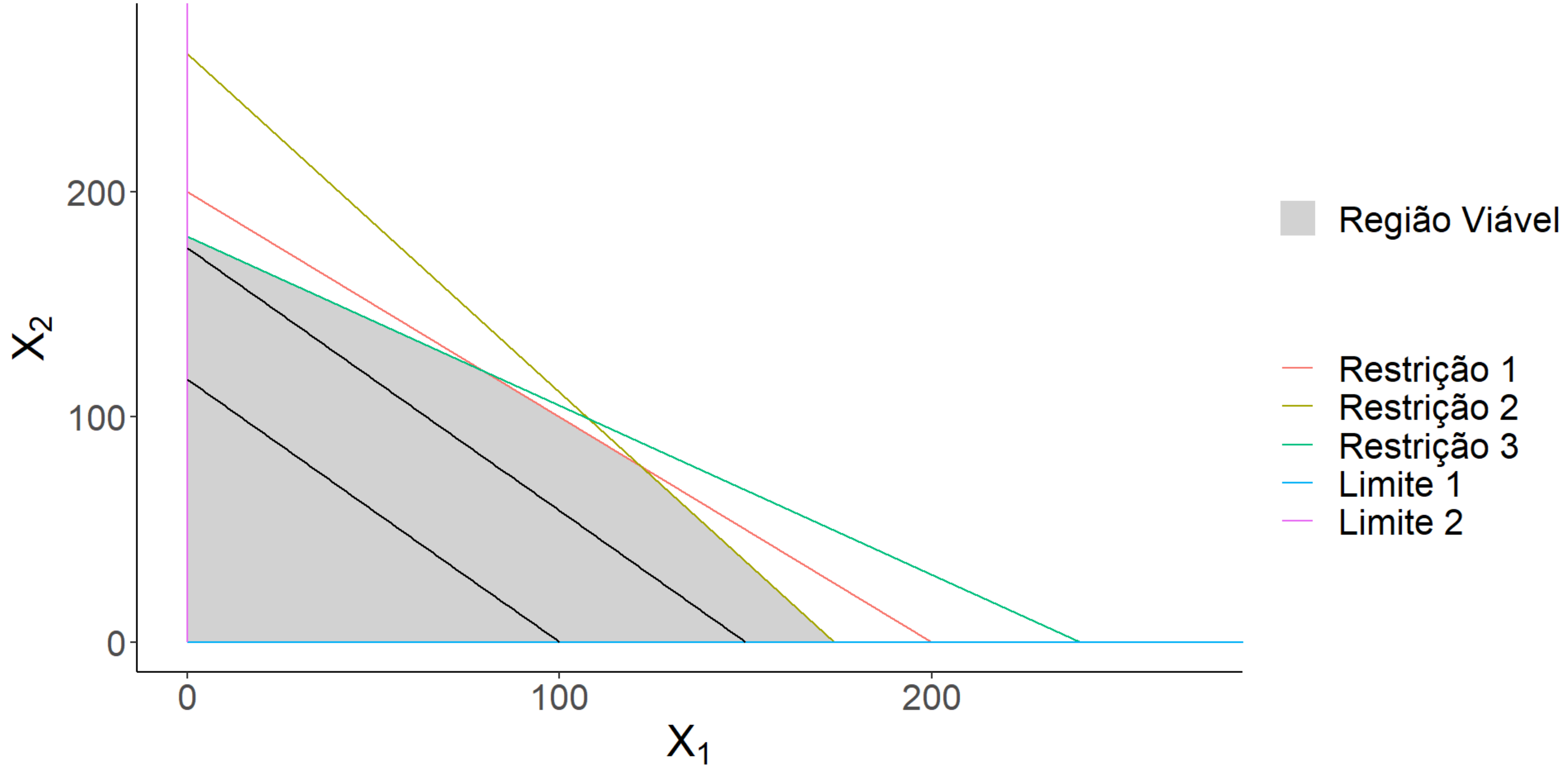
Encontrando a solução do problema graficamente

Função objetivo: $350X_1 + 300X_2 = \mathbf{35000}$



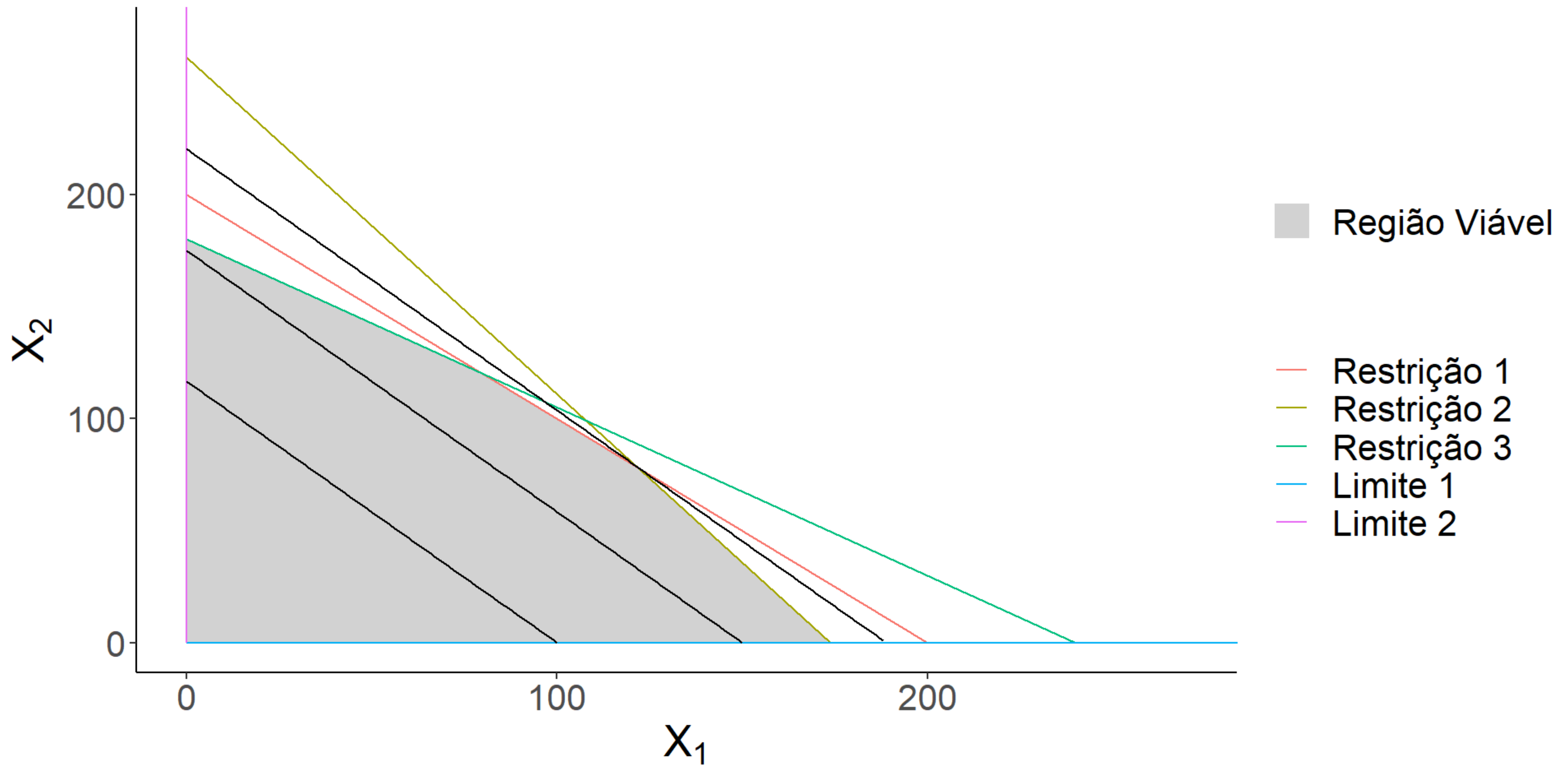
Encontrando a solução do problema graficamente

Função objetivo: $350X_1 + 300X_2 = \mathbf{52500}$



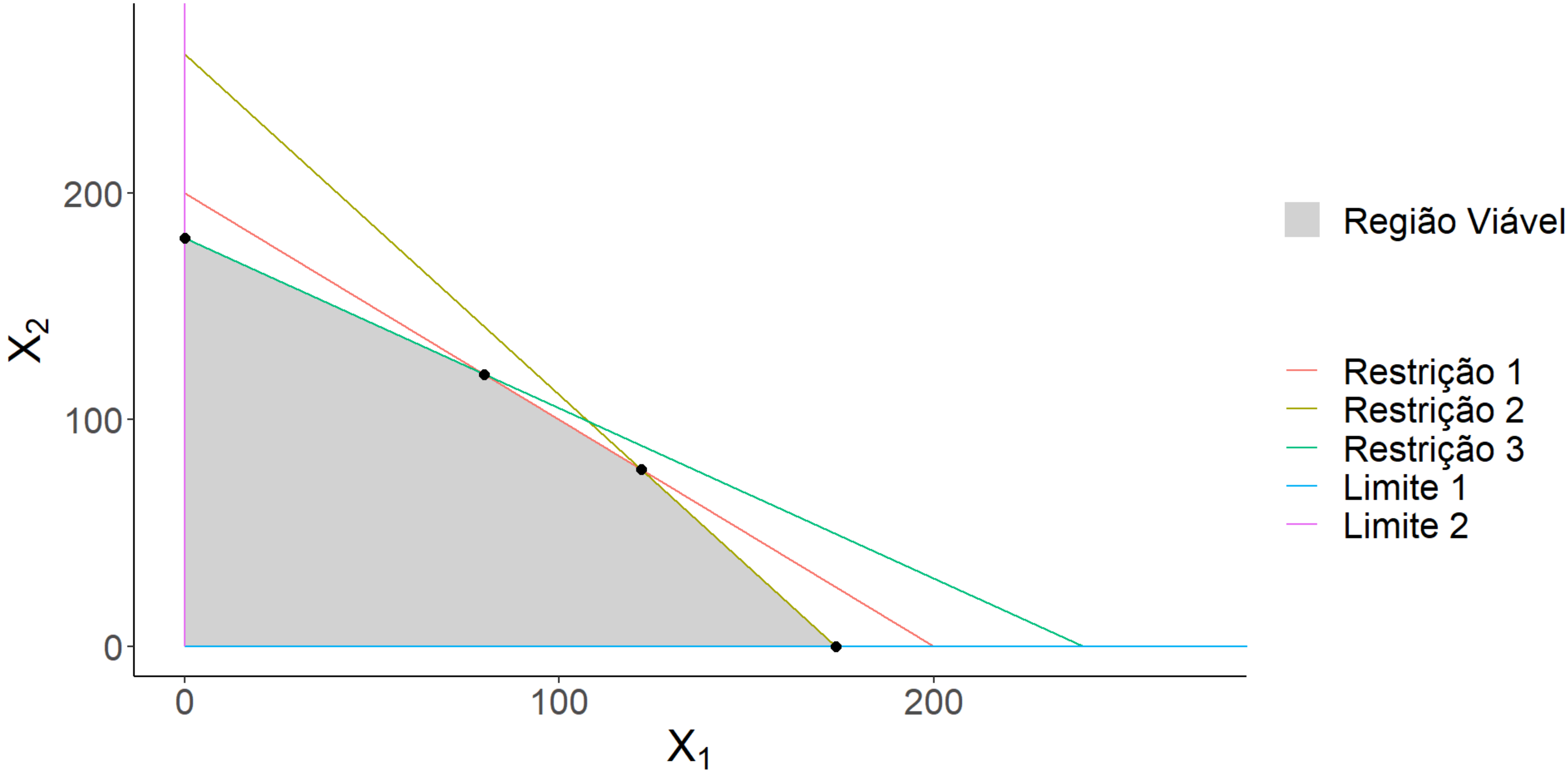
Encontrando a solução do problema graficamente

Função objetivo: $350X_1 + 300X_2 = \mathbf{66100}$



Encontrando a solução pelos pontos de canto

Função objetivo: $350X_1 + 300X_2$



Solução usando R

$$
\begin{aligned}
\text{MAX:} \quad & 350X_1 + 300X_2 \\
\text{Sujeito a:} \quad & 1X_1 + 1X_2 \leq 200 \\
& 9X_1 + 6X_2 \leq 1566 \\
& 12X_1 + 16X_2 \leq 2880 \\
& X_1 \geq 0 \\
& X_2 \geq 0
\end{aligned}
$$

```r
library(lpSolveAPI)

modelo <- make.lp(ncol = 2)
lp.control(lprec = modelo, sense = "max")
set.objfn(lprec = modelo, obj = c(350, 300))

add.constraint(lprec = modelo, xt = c(1, 1),   type = "<=", rhs = 200)
add.constraint(lprec = modelo, xt = c(9, 6),   type = "<=", rhs = 1566)
add.constraint(lprec = modelo, xt = c(12, 16), type = "<=", rhs = 2880)

solve(modelo)
get.objective(modelo)
get.variables(modelo)
```

Exemplo prático

- A AgroPop armazena grandes quantidades de quatro tipos de rações que pode misturar para atender às especificações de determinado cliente.
- A tabela a seguir mostra as quatro rações, as porcentagens de milho, cereais e sais minerais e o custo por Kg para cada tipo.

	Ração 1	**Ração 2**	**Ração 3**	**Ração 4**
Milho	30%	5%	20%	10%
Cereais	10%	30%	15%	10%
Minerais	20%	20%	20%	30%
Custo por Kg	0,25	0,30	0,32	0,15

- A AgroPop acabou de receber um pedido para fornecer 8ton de ração, que contenha pelo menos 20% de milho, 15% de cereais e 15% de sais minerais.
- O que a AgroPop precisa fazer para atender a esse pedido com um custo mínimo?

Exemplo prático – solução

	Ração 1	Ração 2	Ração 3	Ração 4
Milho	30%	5%	20%	10%
Cereais	10%	30%	15%	10%
Minerais	20%	20%	20%	30%
Custo por Kg	0,25	0,30	0,32	0,15

Requisito mínimo: 20% de milho, 15% de cereais e 15% de sais minerais.

$$
\begin{aligned}
\text{MIN:} \quad & 0{,}25X_1 + 0{,}30X_2 + 0{,}32X_3 + 0{,}15X_4 \\
\text{Sujeito a:} \quad & 1X_1 + 1X_2 + 1X_3 + 1X_4 = 8000 \\
& \frac{0{,}30}{8000}X_1 + \frac{0{,}05}{8000}X_2 + \frac{0{,}20}{8000}X_3 + \frac{0{,}10}{8000}X_4 \geq 0{,}20 \\
& \frac{0{,}10}{8000}X_1 + \frac{0{,}30}{8000}X_2 + \frac{0{,}15}{8000}X_3 + \frac{0{,}10}{8000}X_4 \geq 0{,}15 \\
& \frac{0{,}20}{8000}X_1 + \frac{0{,}20}{8000}X_2 + \frac{0{,}20}{8000}X_3 + \frac{0{,}30}{8000}X_4 \leq 0{,}15 \\
& X_1, X_2, X_3, X_4 \geq 0
\end{aligned}
$$

O método Simplex

- Um método para resolução de problemas de programação linear.
- Requer que todas as restrições sejam expressas como igualdades.
- O que fazer com as restrições do tipo "$\leq$" e "$\geq$"?
- Utilizar uma **variável de folga**.

$$a_1X_1 + a_2X_2 \geq b_1$$
$$\hookrightarrow \; a_1X_1 + a_2X_2 + S_1 = b_1$$

Variáveis de folga do Caso 1

$$
\begin{array}{lll}
\text{MAX:} & 350X_1 + 300X_2 & \} \text{ Lucro} \\
\text{Sujeito a:} & 1X_1 + 1X_2 + S_1 = 200 & \} \text{ restrição de bomba} \\
 & 9X_1 + 6X_2 + S_2 = 1566 & \} \text{ restrição de trabalho} \\
 & 12X_1 + 16X_2 + S_3 = 2880 & \} \text{ restrição de tubulação} \\
 & X_1, X_2, S_1, S_2, S_3 \geq 0 & \} \text{ condições de não negatividade}
\end{array}
$$

Soluções básicas viáveis

- Se houver um total de n variáveis em um sistema de m equações, uma estratégia para encontrar uma solução para o sistema de equações é selecionar quaisquer m variáveis e tentar encontrar valores para elas que resolvam o sistema, assumindo que todas as demais variáveis estão configuradas para seus limites inferiores.
- As m variáveis selecionadas para resolver o sistema de equações em um modelo de PL são chamadas **variáveis básicas**, enquanto as variáveis restantes são denominadas **variáveis não básicas**.
- Se puder ser obtida uma solução para o sistema de equações usando dado conjunto de variáveis básicas, essa solução é chamada **solução básica viável** (SBV).
- Cada SBV corresponde a um dos pontos extremos da região viável para o problema de PL.

Método simplex

- Inicialmente identifica qualquer SBV.
- Em seguida, muda para um ponto extremo adjacente se tal movimento melhora o valor da função objetivo
- Quando nenhum ponto extremo adjacente tem um melhor valor de função objetivo, o ponto extremo atual é o ótimo e o método Simplex termina.

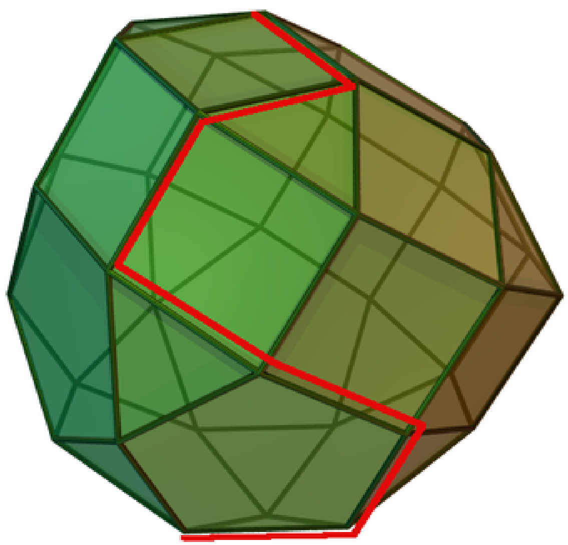


Imagem da Wikipedia

Condições especiais em modelos de PL

- Múltiplas soluções ótimas.
- Restrições redundantes.
- Soluções ilimitadas
- Inviabilidade

Relacionando com aprendizagem estatística

Modelos Preditivos – revisão

- Estudar a relação entre uma variável resposta Y e preditoras $X_1, \ldots, X_p$.
- Relação funcional $Y = f(X)$.
- f é desconhecida
- Objetivo estimar o valor de Y dadas as variáveis preditoras $(X_1, \ldots, X_p)$.
- Estimativa para f: $\hat{f}$.
- Não há interesse particular na forma de $\hat{f}$, desde que ela gere boas predições para Y.
- **Este é o foco da aprendizagem estatística.**

Como estimar f?

- Considere que foi observado um conjunto de dados tamanho n.
- Essas observações são divididas em duas partes: **conjunto de treinamento** e **conjunto de testes** e serão usadas para treinar o modelo para estimar f.
- Seja $x_i = (x_{i1}, \ldots, x_{ip})$ o vetor que presenta os valores das p preditoras na observação i.
- Notação: $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$.
- Ideia: utilizar a informação contida nos dados de treinamento para construir uma função $\hat{f}$, que será nossa estimativa para f, e assim prever $\hat{Y} = \hat{f}(X)$.

Métodos paramétricos

Envolvem uma abordagem baseada em dois passos:

Avaliando a qualidade do ajuste

- Precisamos de uma maneira de medir a performance de um método de aprendizagem estatística.
- Quão boas as predições são em relação aos dados observados?
- No contexto de regressão, a medida mais utilizada é o *erro quadrático médio* (MSE – *mean squared error*):

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left[y_i - \hat{f}\left(x_i\right)\right]^2.$$

- $\hat{f}\left(x_i\right)$ é a predição que $\hat{f}$ dá para a i–ésima observação da amostra.
- *MSE* vai ser pequeno se as predições forem próximas dos valores verdadeiros.
- **MSE de treino**: calculado com as observações do conjunto de treino/treinamento.
- **MSE de validação/teste**: calculado com as observações do conjunto de validação/teste.

Métodos paramétricos

- Em resumo, a abordagem paramétrica reduz o problema de estimação de f para o de estimar um conjunto de parâmetros.
- Potencial desvantagem: o modelo escolhido não representar a forma verdadeira da função f desconhecida. Isso causa estimativas ruins de $\hat{Y}$.
- Possível solução: escolher um modelo *mais flexível*.
- Em geral, quanto mais complexo o modelo, maior o número de parâmetros que precisam ser estimados. Isso pode levar ao problema de *overfitting*.
- Ao fazer uma suposição sobre a forma de f, estamos definindo um **modelo prescritivo**.

Conclusão

- O que é otimizacao matemática.
- Caracterização de um problema de otimização linear.
- Como expressá-lo em temos matemáticos.
- Método Simplex.
- Relacao com aprendizagem estatistica.

Obrigado!

magnotfs@insper.edu.br